



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/601,350	06/23/2003	Jonathan H. Connell	YOR920030166US1	7454

7590 11/29/2006

Ryan, Mason & Lewis, LLP
90 Forest Avenue
Locust Valley, NY 11560

EXAMINER

ARMSTRONG, ANGELA A

ART UNIT PAPER NUMBER

2626

DATE MAILED: 11/29/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary

Application No.

10/601,350

Applicant(s)

CONNELL ET AL.

Examiner

Angela A. Armstrong

Art Unit

2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 11 September 2006.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-22 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-22 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. _____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____ |

DETAILED ACTION

Claim Rejections - 35 USC § 103

The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

1. Claims 1-22 are rejected under 35 U.S.C. 103(a) as being unpatentable over Garg et al, “Frame-dependent multi-stream reliability indicators for audio-visual speech recognition,” Proceedings of International Conference on Acoustics, Speech and Signal Processing, ICASSP 2003, vol. 1, April 2003, pages 24-27 in view of Masai et al (US Patent Application Publication 2003/0177005).
2. Regarding claim 1, Garg teaches a method for audio-visual speech recognition comprising: providing an acoustic-only data model and an acoustic-visual data model (pages 24-26; section 2, entitled “The Multi-Stream HMM”; section 3, entitled “Stream Reliability Indicators”; section 4, entitled “Reliability Based Stream Exponents.”); and decoding at least a portion of an input spoken utterance using selected data models (pages 24-26; section 2, entitled “The Multi-Stream HMM”; section 3, entitled “Stream Reliability Indicators”; section 4, entitled “Reliability Based Stream Exponents”; Tables 1-2). Garg does not specifically teach a data model is selected based on a condition associated with the environment of the speaker. However, selecting an optimum data model for performing recognition based on environmental conditions so as to improve recognition accuracy and performance was well known in the art of speech recognition. Masai discloses (paragraph 75) a method and device for producing acoustic models for recognition and specifically teaches the speech recognition unit recognizes the speech data and convert them into text data in accordance with the environment information of

the time when the speech data are uttered, the acoustic model for recognition selection unit selects the acoustic model for recognition according to the environment information and converts the speech data into text data by using the selected acoustic model for recognition.

It would have been obvious to one of ordinary skill at the time of the invention to modify the system of Garg to allow for the selection of the most optimum data model, as suggested by Masai, for the purpose of improving recognition accuracy and performance of the speech recognizer, as was well known in the art.

Regarding claim 2, Garg and Masai teach storing the acoustic-only data model and the acoustic-visual data model in memory such that model selection is made by shifting one or more pointers to one or more memory locations where the selected model is located (Page 26-27, section 5, "Database and Experiments").

Regarding claim 3, Garg and Masai teach model selection is based on a likelihood ratio test (pages 24-26; section 2, entitled "The Multi-Stream HMM"; section 3, entitled "Stream Reliability Indicators"; section 4, entitled "Reliability Based Stream Exponents").

Regarding claim 4, Garg and Masai teach model selection comprises selecting the acoustic-only data model when a result of the likelihood test is not greater than a threshold value (pages 24-26; section 2, entitled "The Multi-Stream HMM"; section 3, entitled "Stream Reliability Indicators"; section 4, entitled "Reliability Based Stream Exponents").

Regarding claim 5, Garg and Masai teach the model selection step comprises selecting the acoustic-visual data mode when a result of the likelihood test is not less than a threshold (pages 24-26; section 2, entitled "The Multi-Stream HMM"; section 3, entitled "Stream Reliability Indicators"; section 4, entitled "Reliability Based Stream Exponents").

Regarding claim 6, Garg and Masai teach the threshold value is based on a cost associated with a recognition error (Tables 1 and 2; section 3, "Stream Reliability Indicators).

Regarding claim 7, Garg and Masai teach the likelihood ratio test is based on one or more observations of a given visual feature (Tables 1 and 2; section 3, "Stream Reliability Indicators).

Regarding claim 8, Garg and Masai teach the given visual feature is associated with the mouth region of a speaker of the input utterance (Page 26-27, section 5, "Database and Experiments").

Regarding claim 9, Garg and Masai teach the model selection is performed at a rate substantially equivalent to an observation rate associated with the audio-visual speech recognition system (Page 26-27, section 5, "Database and Experiments").

3. Regarding claims 10-22; claims 10-22 are similar in scope and content to method claims 1-9 and are therefore rejected under similar rationale.

Response to Arguments

4. Applicant's arguments filed September 11, 2006, have been fully considered but they are not persuasive. Applicant argues Garg fails to disclose selecting between an acoustic-only data model and an acoustic-visual data model based on a condition associated with a visual environment, and decoding at least a portion of an input spoken utterance using the selected data model and that Masai contains no disclosure relating to a selection between an acoustic-only model and an acoustic-visual model. Applicant further argues neither Garg nor Masai individually teach or suggest the limitations of the independent claims and therefore the

Art Unit: 2626

combination of Garg and Masai also fails to teach or suggest the limitations of the independent claims.

In response to applicant's arguments against the references individually, one cannot show nonobviousness by attacking references individually where the rejections are based on combinations of references. See *In re Keller*, 642 F.2d 413, 208 USPQ 871 (CCPA 1981); *In re Merck & Co.*, 800 F.2d 1091, 231 USPQ 375 (Fed. Cir. 1986). In this instance, Garg was cited for teaching a method for audio-visual speech recognition implementing an acoustic-only data model and an acoustic-visual data model. While, Garg does not specifically teach a data model is selected based on a condition associated with the environment of the speaker, it was well known in the art to provide a means for selecting an optimum data model for performing recognition based on environmental conditions so as to improve recognition accuracy and performance. Masai was cited for teaching this optimum data model selection. Masai discloses a method and device for producing acoustic models for recognition and specifically teaches the speech recognition unit recognizes the speech data and convert them into text data in accordance with the environment information of the time when the speech data are uttered, the acoustic model for recognition selection unit selects the acoustic model for recognition according to the environment information and converts the speech data into text data by using the selected acoustic model for recognition. Thus, the combination of Garg and Masai would provide for a speech recognition system, which utilizes acoustic-only data models and acoustic-visual data models (as provided by Garg), such that the most optimum sets of acoustic only and/or acoustic-visual data models are selected and used for recognition as determined by environment information of the time when the speech data is received (as provided by Masai).

Applicant argues Masai only describes selection of an acoustic data model in accordance with surrounding acoustics, not general environmental conditions. The Examiner cannot concur, and argues that Masai, at least at paragraphs [71 and 72], teaches the environmental conditions can be a time information, a place information, a speaker's physical condition, a conversing partner of the speaker, or data regarding whether the current location is inside the company or inside the home, whether it is during the conference or during the meal. Thus, the teachings of Masai provides adequate support for the limitation and evidence that it is well known in the art to provide a means for selecting an optimum data model for performing recognition based on environmental conditions so as to improve recognition accuracy and performance.

Conclusion

THIS ACTION IS MADE FINAL. Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

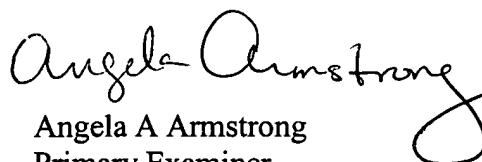
A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Art Unit: 2626

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Angela A. Armstrong whose telephone number is 571-272-7598. The examiner can normally be reached on Monday-Thursday 11:30-8:00 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David Hudspeth can be reached on 571-272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.


Angela A. Armstrong
Primary Examiner
Art Unit 2626

AAA
November 27, 2006